Statistics 541 (updated 2 February)
Spring, 2021

Multivariate Statistics

Professor David W. Scott
2057 Duncan Hall
Ext. 6037, scottdw@rice.edu

Canvas + http://www.stat.rice.edu/∼scottdw/stat541/

Tu/Th 11:20-12:40 HRZ AMP (after 2/16 at least)
Office Hours: By ZOOM (Friday 3-4) or appt., email.

TA: Brian King (bking@rice.edu
Office Hours: By appt/email/zoom
Help Session Wed 6-7 (ZOOM)
Help Sessions Thur 5-6 (ZOOM)

Survey of topics in classical multivariate statistics—roughly equal portions of theory, application, and computing. Additional topics as warranted. You may use the computing language of your choice. (**R** is preferred and will be used in class.) **Mathematica** will be used for symbolic calculations. **JMP** will be used for graphics and **SAS**-like analyses. **Mathematica** and **JMP** are available to Rice students to download on your laptops at no cost (site licenses). We will cover the selections from the book that focus on multivariate normal theory and extensions, the core material on multivariate techniques coming from Chapters 11–20. Special focus on geometric insights.

Grading

- 30% Homework (not pledged, but submit your own paper)
- 35% Exam I (Take home, open book) TBA
- 35% 2-person group data analysis project (depending on class size), oral (in-class, zoom?) presentations TBA; individual written reports due by end of exam period)
- 0% Exam II (optional)

Text: Wolfgang Härdle and Léopold Simar, *Applied Multivariate Statistical Analysis*, Springer, Fifth Edition, 2019.

Optional Text: K.V. Mardia, J.T. Kent, and J.M. Bibby (1979). *Multivariate Analysis*, Academic Press, London (useful but awaiting new edition).

You should plan to review your theoretical linear algebra knowledge (projections, subspaces, matrix operations, bases, linear equations, determinants, eigenvectors, quadratic forms, etc) as we proceed through the course. Gilbert Strang's book is solid. Numerical algorithms are standard these days and will not be discussed in depth.

You should read ahead and ask questions during class. (I had hoped that the in-class experience would facilitate such interactions. We will do our best as long as zoom is the only option.) We will follow the book more closely after the first few weeks.

There are many *R* language references. A good one is *R in a Nutshell: A Desktop Quick Reference* by J. Adler (O'Reilly, $\approx$ \$40). Plenty of free on-line links. You can download R executables for OSX/Linux/Windows from http://cran.r-project.org. Others: RStudio, IDE.

You should be able to download Mathematica and JMP from `it.rice.edu` under the hardware/software tab. (JMP's site license was recently canceled, but may be reinstated for this class.)

Learning Outcomes

- Understand the role of mathematics in probability and statistics
- Perform exploratory data analysis: summarize and visualize data
- Understand the role of the multivariate normal pdf
- Solve problems of mathematical statistics with random samples
- Solve various constrained optimization problems
- Under the classical algorithms of multivariate analysis
- Define, perform and interpret appropriate statistical analyses
- Learn to perform simulations and basic analyses with the statistical software, R
- Communicate statistical ideas and test results

Disability Statement

Any student with a disability requiring accommodations in this course should contact me early in the semester. Additionally, students will also need to contact Disability Support Services in the Ley Student Center.

Honor Code

Please follow all instructions on assignments and exams. Any suspected violations of the Honor Code will be reported to the Honor Council.